ELSEVIER Linear Algebra and its Applications 299 (1999) 101–117

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Self-adjoint operators and pairs of Hermitian forms over the quaternions

Michael Karow

Institut für Dynamische Systeme, Universität Bremen, Fachbereich 3, Postfach 330 440, D-28334 Bremen, Germany

Received 2 February 1998; accepted 25 July 1999

Submitted by H. Schneider

Abstract

We classify self-adjoint operators and pairs of Hermitian forms over the real quaternions by providing canonical matrix representations. In the preliminaries we discuss the Jordan canonical form theorem for quaternionic linear endomorphisms. © 1999 Elsevier Science Inc. All rights reserved.

Keywords: Quaternions; Hermitian forms; Self-adjoint operators

1. Introduction

Classification theorems for pairs of Hermitian, symmetric, skew-Hermitian and skew-symmetric forms over $\mathbb{C}$, $\mathbb{R}$ and other fields are classical and go back to Weierstrass and Kronecker, see [12] for an overview and 225 references. However, pairs of Hermitian forms over the quaternions have been classified just in 1991 by Sergeichuk [11]. He provided a canonical form for such pairs in which Frobenius- and related matrices occur. In the present paper we give an alternative version of the classification theorem. The canonical form we obtain coincides with the well-known one for pairs of Hermitian forms over $\mathbb{C}$, see [5]. We also provide a representation by real matrices.

The paper is subdivided into two parts, the preliminary and the main part. In the preliminaries we remind the reader of some facts of linear algebra over the

E-mail address: karow@math.uni-bremen.de (M. Karow)

PII: S 0 0 2 4 - 3 7 9 5 (9 9) 0 0 1 6 5 - 2

quaternions and give a possibly new proof of the Jordan canonical form theorem for quaternionic linear endomorphisms. Other proofs of this theorem may be found in [7,14]. In the main part of this paper we classify pairs of quaternionic Hermitian forms where one form is assumed to be nondegenerate. In the proof we apply the same methods which are used in [5,10]. These methods yield a classification of self-adjoint operators too.

2. Preliminaries

2.1. Real quaternions

We consider the non-commutative real division algebra of quaternions

$$\mathbb{H} = \{\lambda_1 + \lambda_2\,\mathrm{i} + \lambda_3\,\mathrm{j} + \lambda_4\mathrm{ij} \mid \lambda_k \in \mathbb{R}\},$$

where $\mathrm{i}^2 = \mathrm{j}^2 = -1$, $\mathrm{ij} = -\mathrm{ji}$. The conjugate and the modulus, the real and the imaginary part of $\lambda = \lambda_1 + \lambda_2\,\mathrm{i} + \lambda_3\,\mathrm{j} + \lambda_4\,\mathrm{ij} \in \mathbb{H}$ are defined by

$$\overline{\lambda} := \lambda_1 - \lambda_2\,\mathrm{i} - \lambda_3\,\mathrm{j} - \lambda_4\,\mathrm{ij}, \quad |\lambda| := \sqrt{\lambda\overline{\lambda}} = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2},$$
$$\mathrm{Re}(\lambda) := \lambda_1 = \tfrac{1}{2}(\lambda + \overline{\lambda}), \quad \mathrm{Im}_{\mathbb{H}}(\lambda) := \lambda_2\,\mathrm{i} + \lambda_3\,\mathrm{j} + \lambda_4\,\mathrm{ij} = \tfrac{1}{2}(\lambda - \overline{\lambda}).$$

The multiplicative inverse of $\lambda \in \mathbb{H}^* := \mathbb{H} \setminus \{0\}$ is given by the formula $\lambda^{-1} = \overline{\lambda}/|\lambda|^2$. The conjugation $\lambda \longmapsto \overline{\lambda}$ is an involutory antiautomorphism (involution) of $\mathbb{H}$, i.e.

$$\overline{\lambda\,\mu} = \overline{\mu}\,\overline{\lambda}, \quad \overline{\overline{\lambda}} = \lambda, \quad \overline{\lambda + \mu} = \overline{\lambda} + \overline{\mu} \quad \text{for all } \lambda, \mu \in \mathbb{H}.$$

Quaternions of the form $\lambda = \lambda_1 + \lambda_2\,\mathrm{i}$ where $\lambda_1, \lambda_2 \in \mathbb{R}$ will be identified with the complex numbers. So each $\lambda \in \mathbb{H}$ can uniquely be written in the form $\lambda = \lambda_1 + \lambda_2\,\mathrm{j}$ with $\lambda_1, \lambda_2 \in \mathbb{C}$. We have the commutation rule

$$\lambda\mathrm{j} = \mathrm{j}\overline{\lambda} \quad \text{for all } \lambda \in \mathbb{C}. \tag{1}$$

The multiplicative group $\mathbb{H}^*$ acts on $\mathbb{H}$ via

$$(\lambda, \mu) \longmapsto a_\mu(\lambda) := \mu^{-1}\lambda\mu \quad \text{for all } \lambda \in \mathbb{H},\ \mu \in \mathbb{H}^*.$$

Two quaternions are called similar if they are in the same orbit of this action. The mappings a_μ are the only $\mathbb{R}$-algebra-automorphisms of $\mathbb{H}$ (see [8]), and we have the equivalence

$$a_\mu = \mathrm{i}d_{\mathbb{H}} \;\Leftrightarrow\; \mu \in \text{center of } \mathbb{H} \;\Leftrightarrow\; \mu \in \mathbb{R}.$$

Proposition 2.1. $\lambda_1, \lambda_2 \in \mathbb{H}$ *are similar iff* $\mathrm{Re}(\lambda_1) = \mathrm{Re}(\lambda_2)$ *and* $|\mathrm{Im}_{\mathbb{H}}(\lambda_1)| = |\mathrm{Im}_{\mathbb{H}}(\lambda_2)|$.

Proof. See [16]. $\square$

Corollary 2.2. *Each similarity class* $c_\lambda := \{a_\mu(\lambda) \mid \mu \in \mathbb{H}^*\}$ *contains exactly one element* $\lambda^\uparrow$ *in the closed upper half-plane* $\mathbb{C}^\uparrow := \{z \in \mathbb{C} \mid \mathrm{Im}(z) \geqslant 0\}$ *of* $\mathbb{C}$. c_λ *contains also the conjugate* $\overline{\lambda^\uparrow}$. c_λ *has only one element iff* $\lambda = \lambda^\uparrow \in \mathbb{R}$.

2.2. Linear mappings and eigenvalues

Let X, Y be right vector spaces over $\mathbb{H}$. A mapping $\phi : X \to Y$ is called semilinear with respect to the automorphism a_μ if

(a) $\phi(x_1 + x_2) = \phi(x_1) + \phi(x_2)$ for all $x_1, x_2 \in X$,

(b) $\phi(x\lambda) = \phi(x)\, a_\mu(\lambda)$ for all $x \in X,\ \lambda \in \mathbb{H}$.

ϕ is called $\mathbb{H}$ (resp. $\mathbb{C}$)-linear if (a) and (b′) hold, where

(b′) $\phi(x\lambda) = \phi(x)\,\lambda$ for all $x \in X, \lambda \in \mathbb{H}$ (resp. $\lambda \in \mathbb{C}$).

Analogous definitions hold for left vector spaces. Since $\mathbb{H}$ is a non-commutative ring the scalar multiplications

$$m_\lambda : X \to X, \quad m_\lambda(x) := x\lambda \quad \text{for all } x \in X,\ \lambda \in \mathbb{H}$$

are not linear in general. However, we have the easily verified proposition.

Proposition 2.3.
(1) m_λ *is* a_λ*-semilinear for all* $\lambda \in \mathbb{H}^*$.
(2) m_λ *is* $\mathbb{H}$*-linear iff* $\lambda \in \mathbb{R}$.
(3) m_λ *is* $\mathbb{C}$*-linear iff* $\lambda \in \mathbb{C}$.

We have to be careful with the notion of linearity if we use matrices, as the following proposition shows.

Proposition 2.4. *For* $A \in \mathbb{H}^{m\times n}$ *let* $\phi_A : \mathbb{H}^n \to \mathbb{H}^m$ *be defined by* $\phi_A(x) := Ax$ *for all* $x \in \mathbb{H}^n$.
(a) *If we regard* $\mathbb{H}^n$ *and* $\mathbb{H}^m$ *as right vector spaces over* $\mathbb{H}$ *then* ϕ_A *is* $\mathbb{H}$*-linear for all* $A \in \mathbb{H}^{m\times n}$. *Moreover, for each* $\mathbb{H}$*-linear map* $\psi : \mathbb{H}^n \to \mathbb{H}^m$ *there exists exactly one matrix* $A \in \mathbb{H}^{m\times n}$ *such that* $\psi = \phi_A$.
(b) *If we regard* $\mathbb{H}^n$ *and* $\mathbb{H}^m$ *as left vector spaces over* $\mathbb{H}$ *then* ϕ_A *is* $\mathbb{H}$*-linear if and only if* $A \in \mathbb{R}^{n\times m}$.

Proof. (a) is obvious. (b) If $A\lambda x = \lambda Ax$ for all $\lambda \in \mathbb{H}$, $x \in \mathbb{H}^n$ then $A\lambda = \lambda A$ for all $\lambda \in \mathbb{H}$, i.e. the entries of A commute with all quaternions and are therefore real. □

In view of this proposition we will use *right* vector spaces throughout this paper. [1] By $\mathrm{End}_{\mathbb{H}}(X)$ we denote the set of all $\mathbb{H}$-linear endomorphisms of the right

[1] Note that each left vector space X over $\mathbb{H}$ can be endowed with the structure of a right vector space by defining multiplication from the right as $x\lambda := \overline{\lambda}x$ for all $x \in X,\ \lambda \in \mathbb{H}$.

vector space X. X is supposed to be finite-dimensional from now on. A basis $\mathfrak{B} = (b_1, \ldots, b_n)$ of X induces a coordinate map

$$x = \sum_{k=1}^{n} b_k \, \xi_k \in X \;\longmapsto\; [x]_{\mathfrak{B}} := (\xi_1, \ldots, \xi_n)^{\mathrm{T}} \in \mathbb{H}^n,$$

where T denotes the transpose. For every $\phi \in \mathrm{End}_{\mathbb{H}}(X)$ there exists exactly one matrix $[\phi]_{\mathfrak{B}} \in \mathbb{H}^{n\times n}$ such that

$$[\phi(x)]_{\mathfrak{B}} = [\phi]_{\mathfrak{B}} \, [x]_{\mathfrak{B}} \quad \text{for all } \; x \in X.$$

$\lambda \in \mathbb{H}$ is called an eigenvalue of $\phi \in \mathrm{End}_{\mathbb{H}}(X)$ if there is an $x \in X \setminus \{0\}$ such that $\phi(x) = x\lambda$. Let $\sigma(\phi)$ denote the spectrum of ϕ, i.e. the set of all eigenvalues of ϕ. From $\phi(x) = x\lambda$ it follows that $\phi(x\mu) = x\lambda\mu = (x\mu)(\mu^{-1}\lambda\mu) = (x\mu)a_\mu(\lambda)$ for all $\mu \in \mathbb{H}^*$. So we have

$$\lambda \in \sigma(\phi) \;\Leftrightarrow\; a_\mu(\lambda) \in \sigma(\phi) \;\text{ for all }\; \mu \in \mathbb{H}^* \;\Leftrightarrow\; c_\lambda \subseteq \sigma(\phi). \tag{2}$$

The quaternionic vector space X is a complex vector space of double dimension. If we regard ϕ as a $\mathbb{C}$-linear endomorphism of X, then ϕ has the complex spectrum $\sigma_{\mathbb{C}}(\phi) := \sigma(\phi) \cap \mathbb{C}$. By combining (2) and Corollary 2.2 we obtain the following theorem.

Theorem 2.5. *Let $\phi \in \mathrm{End}_{\mathbb{H}}(X)$. Then $\sigma_{\mathbb{C}}(\phi)$ is symmetric with respect to the real axis and the spectrum of ϕ is*

$$\sigma(\phi) = \biguplus_{\lambda \in \sigma(\phi)\cap\mathbb{C}^{\uparrow}} c_\lambda,$$

where $\uplus$ denotes the disjoint union.

For quaternionic square matrices two kinds of eigenvalues can be defined: $\lambda \in \mathbb{H}$ is called a left (right) eigenvalue of $A \in \mathbb{H}^{n\times n}$ if $Ax = \lambda\, x$ ($Ax = x\lambda$) for some $x \in \mathbb{H}^n \setminus \{0\}$. Both kinds of eigenvalues have been investigated in the literature, see e.g. [2,4,9,15,16]. The right eigenvalues are exactly the eigenvalues of $\phi_A \in \mathrm{End}_{\mathbb{H}}(\mathbb{H}^n)$, where ϕ_A is defined as in Proposition 2.4. The left eigenvalues do not fit into our abstract setup and will not be used in this paper.

2.3. The Jordan canonical form

We will discuss now the Jordan canonical form for linear endomorphisms of quaternionic right vector spaces. To this end we introduce some notation: $\oplus$ denotes the direct sum of matrices and the direct sum of quaternionic subspaces, while $\oplus_{\mathbb{C}}$ denotes the direct sum of complex subspaces. For $\phi \in \mathrm{End}_{\mathbb{H}}(X)$ the ϕ-invariant subspace generated by $x \in X$ is given by

$$\langle x \rangle_\phi := \left\{ \sum_{k=0}^{n} \phi^k(x)\lambda_k \;\middle|\; n \in \mathbb{N},\; \lambda_k \in \mathbb{H} \right\}.$$

An upper triangular Jordan block is denoted by

$$\mathscr{J}_p(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \in \mathbb{H}^{p\times p}.$$

In Section 3 we will also use the real Jordan blocks

$$\mathscr{J}_{2p}(a,b) := \begin{bmatrix} \Lambda & I & & 0 \\ & \Lambda & \ddots & \\ & & \ddots & I \\ 0 & & & \Lambda \end{bmatrix} \in \mathbb{R}^{2p\times 2p},$$

$$\Lambda := \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

By a system of vectors we mean an n-tuple of vectors in a vector space X.

The Jordan decomposition theorem for nilpotent endomorphisms is as follows.

Theorem 2.6. *Let $\phi \in \mathrm{End}_{\mathbb{H}}(X)$ be nilpotent (i.e. $\phi^p = 0$ for some $p \in \mathbb{N}_{\geqslant 1} = \{1, 2, \ldots\}$). Then there exist vectors $x_{p,k} \in X\,(p \in \mathbb{N}_{\geqslant 1}, 1 \leqslant k \leqslant d_p \in \mathbb{N} = \{0, 1, \ldots\})$ such that*

$$X = \bigoplus_{p\in\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p} \langle x_{p,k}\rangle_\phi \quad \text{and} \quad \phi^p(x_{p,k}) = 0 \neq \phi^{p-1}(x_{p,k}) \ \text{for all}\, p, k.$$

The system

$$\begin{aligned}\mathfrak{B} := (&\phi^{p_0-1}(x_{p_0,1}), \ldots, \phi^{p-1}(x_{p,k}), \phi^{p-2}(x_{p,k}), \ldots, \\ &\phi(x_{p,k}), x_{p,k}, \ldots, x_{p_1,d_{p_1}}) \\ (p_0 &:= \min\{p \,|\, d_p \neq 0\}, \ \ p_1 := \max\{p \mid d_p \neq 0\})\end{aligned}$$

is a basis of X and with respect to this basis ϕ has the matrix representation

$$[\phi]_{\mathfrak{B}} = \bigoplus_{p\,\in\,\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p} \mathscr{J}_p(0).$$

In such a decomposition the numbers d_p are uniquely determined and given as follows. Let $U_p := \ker\phi^{p-1} + (\ker\phi^p \cap \mathrm{range}\ \phi)$. If $U_p = \ker\phi^p$ then $d_p = 0$. If $U_p \neq \ker\phi^p$ then $d_p \neq 0$ and the system $(x_{p,k} + U_p)_{1\leqslant k\leqslant d_p}$ is a basis of the quotient space $\ker\phi^p/U_p$. Thus

$$d_p = \dim(\ker\phi^p/U_p) \quad \text{for all } p \in \mathbb{N}_{\geqslant 1}.$$

Proof. The standard proofs of this theorem for nilpotent endomorphisms over fields work for skew fields too; especially for $\mathbb{H}$. □

For each $\phi \in \mathrm{End}_{\mathbb{H}}(X)$ and all $\lambda \in \mathbb{C}$ let

$$E_\phi(\lambda) := \{x \in X \mid (\phi - m_\lambda)^p(x) = 0 \text{ for some } p \in \mathbb{N}_{\geqslant 1}\}.$$

If we regard X as vector space over $\mathbb{C}$ and ϕ as $\mathbb{C}$-linear endomorphism of X then $E_\phi(\lambda)$ is the generalized eigenspace of ϕ with respect to the eigenvalue λ. Especially, $E_\phi(\lambda) \neq \{0\}$ iff $\lambda \in \sigma(\phi)$.

Theorem 2.7 (Jordan decomposition theorem). *Let X be a finite-dimensional right vector space over $\mathbb{H}$ and $\phi \in \mathrm{End}_{\mathbb{H}}(X)$.*

(a) *If $\lambda \in \mathbb{R}$, then $E_\phi(\lambda)$ is a ϕ-invariant quaternionic subspace of X.*

(b) *If $\lambda \in \mathbb{C} \setminus \mathbb{R}$ then $E_\phi(\lambda)$ is a ϕ-invariant complex subspace of X and the multiplication by $\jmath$ is a $\mathbb{C}$-antilinear isomorphism*

$$m_\jmath|_{E_\phi(\lambda)} : E_\phi(\lambda) \to E_\phi(\bar{\lambda});$$

especially $\dim_{\mathbb{C}} E_\phi(\lambda) = \dim_{\mathbb{C}} E_\phi(\bar{\lambda})$. *The direct sum $E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\bar{\lambda})$ is a ϕ-invariant quaternionic subspace of X and each basis of the complex space $E_\phi(\lambda)$ is also a basis of the quaternionic space $E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\bar{\lambda})$.*

(c) *X has the direct decomposition*

$$X = \bigoplus_{\lambda \in \sigma(\phi) \cap \mathbb{R}} E_\phi(\lambda) \oplus \bigoplus_{\lambda \in (\sigma(\phi) \cap \mathbb{C}^{\uparrow}) \setminus \mathbb{R}} \left(E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\bar{\lambda})\right). \tag{3}$$

(d) *ϕ has a matrix representation*

$$[\phi]_{\mathfrak{B}} = \bigoplus_{\lambda \in \sigma(\phi) \cap \mathbb{C}^{\uparrow}} \bigoplus_{p \in \mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \mathscr{J}_p(\lambda).$$

The numbers $d_p(\phi, \lambda)$ are unique and by choice of a suitable basis $\widetilde{\mathfrak{B}}$ instead of $\mathfrak{B}$ each Jordan block $\mathscr{J}_p(\lambda)$ can be replaced by $\mathscr{J}_p(\tilde{\lambda})$, where $\tilde{\lambda} \in c_\lambda$.

Proof. (a) holds since $\phi - m_\lambda$ is $\mathbb{H}$-linear if $\lambda \in \mathbb{R}$. The statements of (b) except for the last follow from the commutation rule (1). To show the last statement, let $\mathfrak{S} = (x_1, \ldots, x_r)$ be a $\mathbb{C}$-linearly independent system of vectors in $E_\phi(\lambda)$ and $\mu_k^{(1)}$, $\mu_k^{(2)} \in \mathbb{C}$, $k \in \{1, \ldots, r\}$, such that $0 = \sum_{k=1}^r x_k(\mu_k^{(1)} + \mu_k^{(2)} \jmath) = y_1 + y_2 \jmath$, where $y_l := \sum_{k=1}^r x_k \mu_k^{(l)}$, $l = 1, 2$. Then we have $y_1, y_2 \in E_\phi(\lambda)$ and $y_2 \jmath \in E_\phi(\bar{\lambda})$. From $E_\phi(\lambda) \cap E_\phi(\bar{\lambda}) = \{0\}$ it follows that $y_1 = y_2 = 0$, and from the $\mathbb{C}$-linear independence of $\mathfrak{S}$ we conclude that $\mu_k^{(1)} = \mu_k^{(2)} = 0$. So $\mathfrak{S}$ is linearly independent over $\mathbb{H}$ and if $\mathfrak{S}$ is a basis of the complex vector space $E_\phi(\lambda)$ then $\mathfrak{S}$ is a basis of the quaternionic space $E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\bar{\lambda})$. (c) is a consequence of (a), (b) and the Jordan decomposition theorem for $\mathbb{C}$-linear endomorphisms. (d) Because of (c) we can treat the quaternionic subspaces in the decomposition (3) separately. Consider first a $\lambda \in \sigma(\phi) \cap \mathbb{R}$. Then $\psi := (\phi - m_\lambda)|_{E_\phi(\lambda)}$ is a nilpotent $\mathbb{H}$-linear endomorphism of $E_\phi(\lambda)$ and according to the last theorem there is a basis $\mathfrak{B}_\lambda$ of $E_\phi(\lambda)$ such that

$$[\psi]_{\mathfrak{B}_\lambda} = \bigoplus_{p \in \mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \mathscr{J}_p(0)$$

with uniquely determined $d_p(\phi, \lambda)$. Since $[\phi|_{E_\phi(\lambda)}]_{\mathfrak{B}_\lambda} = [m_\lambda|_{E_\phi(\lambda)} + \psi]_{\mathfrak{B}_\lambda} = [m_\lambda]_{\mathfrak{B}_\lambda} + [\psi]_{\mathfrak{B}_\lambda}$ we have

$$[\phi|_{E_\phi(\lambda)}]_{\mathfrak{B}_\lambda} = \bigoplus_{p \in \mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \mathscr{J}_p(\lambda). \tag{4}$$

Next take a $\lambda \in (\sigma(\phi) \cap \mathbb{C}^{\uparrow}) \setminus \mathbb{R}$. By the Jordan decomposition theorem for $\mathbb{C}$-linear endomorphisms there is a basis $\mathfrak{B}_\lambda$ of the complex space $E_\phi(\lambda)$ such that (4) holds. However, by (b), $\mathfrak{B}_\lambda$ is a quaternionic basis of $E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\overline{\lambda})$, so, if we regard $\phi|_{E_\phi(\lambda)\oplus_{\mathbb{C}} E_\phi(\overline{\lambda})}$ as a quaternionic endomorphism,

$$[\phi|_{E_\phi(\lambda)\oplus_{\mathbb{C}} E_\phi(\overline{\lambda})}]_{\mathfrak{B}_\lambda} = \bigoplus_{p \in \mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \mathscr{J}_p(\lambda).$$

Finally let $(x_1, \ldots, x_p)$ be a Jordan chain for ϕ in $\mathfrak{B}_\lambda$ corresponding to the Jordan block $\mathscr{J}_p(\lambda)$, i.e.

$$\phi(x_k) = x_k\,\lambda + x_{k-1} \quad \text{for } 1 \leqslant k \leqslant p \text{ where } x_0 := 0.$$

If $\widetilde{\lambda} = \mu^{-1}\lambda\mu$ with $\mu \in \mathbb{H}^*$ then we have

$$\phi(x_k\,\mu) = (x_k\,\mu)\,\widetilde{\lambda} + (x_{k-1}\,\mu) \quad \text{for } 1 \leqslant k \leqslant p,$$

so $(x_1\,\mu, \ldots, x_p\mu)$ is a Jordan chain corresponding to the block $\mathscr{J}_p(\widetilde{\lambda})$. $\square$

2.4. Hermitian forms

Definition 2.8. A *Hermitian form* on a right vector space X over $\mathbb{H}$ is a mapping $h : X \times X \to \mathbb{H}$ with the following properties:

(a) $h(x_1 + y_1, x_2 + y_2) = h(x_1, x_2) + h(x_1, y_2) + h(y_1, x_2) + h(y_1, y_2)$ for all $x_1, x_2, y_1, y_2 \in X$;

(b) $h(x\lambda, y) = \overline{\lambda} h(x, y)$, $\quad h(x, y\lambda) = h(x, y)\lambda$ for all $x, y \in X$, $\lambda \in \mathbb{H}$;

(c) $h(y, x) = \overline{h(x, y)}$.

From (c) it follows that

$$h(x, x) \in \mathbb{R} \quad \text{for all } x \in X. \tag{5}$$

Moreover the following so called *polarization identity* holds.

$$\begin{aligned} 4h(x,y) = {} & h(x+y, x+y) - h(x-y, x-y) \\ & + \mathrm{i}\,(h(x\mathrm{i}+y, x\mathrm{i}+y) - h(x\mathrm{i}-y, x\mathrm{i}-y)) \\ & + \mathrm{j}\,(h(x\mathrm{j}+y, x\mathrm{j}+y) - h(x\mathrm{j}-y, x\mathrm{j}-y)) \\ & + \mathrm{ij}\,(h(x\mathrm{ij}+y, x\mathrm{ij}+y) - h(x\mathrm{ij}-y, x\mathrm{ij}-y)) \quad \text{for all } x, y \in X. \end{aligned}$$

Consequently,

$$h = 0 \quad \text{iff} \quad h(x,x) = 0 \ \text{ for all } x \in X. \tag{6}$$

A Hermitian form h is called nondegenerate if the antilinear mapping

$$\widehat{h} : X \to X^*, \quad \widehat{h}(x)(y) := h(x,y) \quad \text{for all } x, y \in X$$

is bijective, where X^* is the dual space of X, i.e. the left vector space of all $\mathbb{H}$-linear mappings $f : X \to \mathbb{H}$. Two vectors $x, y \in X$ are called *orthogonal* if $h(x,y) = 0$. The *orthogonal companion* of a subspace $U \subseteq X$ is defined as

$$U^{\perp_h} := \{x \in X \mid h(x,y) = 0 \ \text{ for all } y \in U\}.$$

If the restriction $h|_{U \times U}$ is nondegenerate, then we have the direct decomposition

$$X = U \oplus U^{\perp_h}.$$

See [3] for a proof. A matrix $A \in \mathbb{H}^{r \times r}$ is called Hermitian if $A^* = A$, where A^* denotes the conjugate tranpose of A. Let $\mathfrak{S} = (x_1, \ldots, x_r)$ be a system of vectors in X. Then the Hermitian matrix $[h]_{\mathfrak{S}} := (h(x_j, x_k))_{j,k \leqslant r}$ is called the *Gramian* of $\mathfrak{S}$ with respect to h. We will need the following lemma.

Lemma 2.9. *If $[h]_{\mathfrak{S}}$ is nonsingular, then $\mathfrak{S}$ is linearly independent.*

Proof. From $\sum_{k=1}^{r} x_k \lambda_k = 0$, it follows that $\sum_{k=1}^{r} h(x_j, x_k)\lambda_k = 0$ for all $j \leqslant r$, in matrix notation $[h]_{\mathfrak{S}}(\lambda_1, \ldots \lambda_r)^{\mathrm{T}} = 0$, and so $\lambda_k = 0$ for all $k \leqslant r$, because $[h]_{\mathfrak{S}}$ is nonsingular. □

Let $\mathfrak{B} = (b_1, \ldots, b_n)$, $\mathfrak{C} = (c_1, \ldots, c_n)$ be bases of X and let $S \in \mathbb{H}^{n \times n}$ be the matrix with $[x]_{\mathfrak{B}} = S[x]_{\mathfrak{C}}$ for all $x \in X$. Then

- $h(x,y) = [x]_{\mathfrak{B}}^* \, [h]_{\mathfrak{B}} \, [y]_{\mathfrak{B}}$ for all $x, y \in X$,
- h is nondegenerate iff $[h]_{\mathfrak{B}}$ is invertible,
- $[h]_{\mathfrak{C}} = S^* \, [h]_{\mathfrak{B}} \, S$.

Hermitian forms on finite-dimensional quaternionic spaces are completely classified by the following.

Theorem 2.10 (Sylvester's Law of Inertia). *Let h be a Hermitian form on the finite-dimensional quaternionic vector space X. Then X has a basis $\mathfrak{B}$ such that*

$$[h]_{\mathfrak{B}} = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_r) \oplus 0_p,$$

where 0_p *is the* $p \times p$ *zero-matrix and* $\varepsilon_k \in \{-1, 1\}$. *In such a representation the numbers* r, $\pi(h) := \sharp\{k \mid \varepsilon_k = 1\}$ *and* $\nu(h) := \sharp\{k \mid \varepsilon_k = -1\}$ *are uniquely determined*, *namely*

$$\begin{aligned} r &= \operatorname{rank} \widehat{h}, \\ \pi(h) &= \max\{\dim U \mid U \textit{ subspace of } X,\ h(x,x) > 0 \ \textit{for all } x \in U \setminus \{0\}\}, \\ \nu(h) &= \max\{\dim U \mid U \textit{ subspace of } X,\ h(x,x) < 0 \ \textit{for all } x \in U \setminus \{0\}\}. \end{aligned}$$

We define

$$\operatorname{signature}(h) := \pi(h) - \nu(h) = \sum_{k=1}^{r} \varepsilon_k.$$

Proof. See [3]. □

2.5. *Self-adjoint operators*

Definition 2.11. An operator $\phi \in \operatorname{End}_{\mathbb{H}}(X)$ is called self-adjoint with respect to a nondegenerate Hermitian form h on X, if

$$h(x, \phi(y)) = h(\phi(x), y) \quad \text{for all } x, y \in X.$$

The basic facts about self-adjoint operators and their connection to Hermitian forms are given in the following proposition.

Proposition 2.12. *Let* $h_0 : X \times X \to \mathbb{H}$ *be nondegenerate*, $\dim X \in \mathbb{N}_{\geqslant 1}$.

1. *For every Hermitian form* h_1 *on* X *there is exactly one* $\phi \in \operatorname{End}_{\mathbb{H}}(X)$ *such that*

   $$h_1(x, y) = h_0(x, \phi(y)) \quad \textit{for all } x, y \in X.$$

   ϕ *is self-adjoint with respect to* h_0.
2. *On the other hand let* ϕ *be* h_0*-self-adjoint. Then the mapping*

   $$(x, y) \longmapsto h_0(x, \phi(y)) \quad \textit{for all } x, y \in X$$

   is a Hermitian form on X.
3. *Let* h_1 *be a Hermitian form on X,* $\phi \in \operatorname{End}_{\mathbb{H}}(X)$ *and* $\mathfrak{B}$ *be a basis of X,* $H_0 := [h_0]_{\mathfrak{B}}$, $H_1 := [h_1]_{\mathfrak{B}}$ *and* $F := [\phi]_{\mathfrak{B}}$. *Then*
   (a) ϕ *is* h_0*-self-adjoint iff* $F^* H_0 = H_0 F$;
   (b) $h_1(\cdot, \cdot) = h_0(\cdot, \phi(\cdot))$ *iff* $H_1 = H_0 F$.

Proof. Standard. □

3. Classification of self-adjoint operators and Hermitian pairs

We are now in the position to study our main problem, the classification of self-adjoint operators and Hermitian pairs on a finite-dimensional quaternionic right vector space X. For brevity let us call (h_0, ϕ, h_1) an hs-triple (Hermitian-self-adjoint-triple) on X if $h_0, h_1 : X \times X \to \mathbb{H}$ are Hermitian forms, h_0 is nondegenerate, $\phi \in \mathrm{End}_{\mathbb{H}}(X)$ is h_0-self-adjoint and $h_1(\cdot,\cdot) = h_0(\cdot, \phi(\cdot))$. In the proof of our main result we will need the following lemma.

Lemma 3.1 (Splitting lemma). *Let (h_0, ϕ, h_1) be an hs-triple on X and $\mathfrak{S} = (x_1, \ldots, x_p)$ a system of vectors in X, such that $U := \operatorname{span}\mathfrak{S}$ is ϕ-invariant and $[h_0]_{\mathfrak{S}}$ is nonsingular. Then $W := U^{\perp_{h_0}}$ is also ϕ-invariant, h_0 is nondegenerate on U and W, and we have the direct decomposition $X = U \oplus W$, where $h_1(U, W) = h_0(U, W) = \{0\}$. Moreover $\mathfrak{S}$ is a basis of U, and if $\mathfrak{C} = (c_1, \ldots, c_s)$ is an arbitrary basis of W, then it holds for the basis $\mathfrak{B} := (x_1, \ldots x_r, c_1, \ldots, c_s)$ of X that*

$$([h_0]_{\mathfrak{B}}, [\phi]_{\mathfrak{B}}, [h_1]_{\mathfrak{B}}) = ([h_0]_{\mathfrak{S}},\ [\phi]_{\mathfrak{S}},\ [h_0]_{\mathfrak{S}}[\phi]_{\mathfrak{S}}) \oplus ([h_0]_{\mathfrak{C}},\ [\phi]_{\mathfrak{C}},\ [h_1]_{\mathfrak{C}}),$$

where $\oplus$ denotes the direct sum of matrix triples, i.e.

$$(A_1, B_1, C_1) \oplus (A_2, B_2, C_2) := (A_1 \oplus A_2, B_1 \oplus B_2, C_1 \oplus C_2).$$

Proof. By the ϕ-invariance of U we have

$$h_1(U, W) = h_0(U, \phi(W)) = h_0(\phi(U), W) \subseteq h_0(U, W) = \{0\}.$$

So U,W are orthogonal with respect to h_1 and W is ϕ-invariant. The other facts, except for the nondegeneracy of $h_0|_{W\times W}$, follow from the preliminaries. However, if h_0 were degenerate on W, there were an $x \in W$ such that $\{0\} = h_0(x, W) = h_0(x, U \oplus W) = h_0(x, X)$, so h_0 were degenerate on X. $\square$

We will use this lemma to give matrix representations of hs-triples which are as simple as possible. The h_0-Gramians in these representations are direct sums of scalar multiples of the backward identities

$$\mathscr{E}_p := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} \in \mathbb{H}^{p\times p}, \quad p \in \mathbb{N}_{\geqslant 1},$$

while the h_1-Gramians are built up from blocks of the form

$$\mathscr{E}_p\,\mathscr{J}_p(\lambda) = \begin{bmatrix} 0 & & & \lambda \\ & & & 1 \\ & & \iddots & \\ & \iddots & \iddots & \\ \lambda & 1 & & 0 \end{bmatrix}, \quad \mathscr{E}_{2p}\mathscr{J}_{2p}(a,b) = \begin{bmatrix} 0 & & & M \\ & & & E \\ & & \iddots & \\ & \iddots & \iddots & \\ M & E & & 0 \end{bmatrix},$$

where

$$M := \begin{bmatrix} b & a \\ a & -b \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad E := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We treat the nilpotent case first. $\delta_{j,k}$ denotes the Kronecker symbol.

Theorem 3.2 (The nilpotent case). *Let (h_0, ϕ, h_1) be an hs-triple on X, where ϕ is nilpotent. Then there are vectors $x_{p,k} \in X$ and signs $\varepsilon_k^{(p)} \in \{-1, 1\}$ with*

(a) $X = \bigoplus_{p\in\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p} \langle x_{p,k}\rangle_\phi$ *and* $\phi^p(x_{p,k}) = 0 \neq \phi^{p-1}(x_{p,k})$ *for all* p, k;

(b) $h_0\left(\langle x_{p_1,k_1}\rangle_\phi, \langle x_{p_2,k_2}\rangle_\phi\right) = h_1\left(\langle x_{p_1,k_1}\rangle_\phi, \langle x_{p_2,k_2}\rangle_\phi\right) = 0$ *for* $(p_1, k_1) \neq (p_2, k_2)$;

(c) $h_0\left(x_{p,k}, \phi^j(x_{p,k})\right) = \varepsilon_k^{(p)}\delta_{p-1,j}$ *for all* $j \in \mathbb{N}$.

The matrix representations of h_0, h_1 and ϕ with respect to the basis

$$\mathfrak{B} := (\phi^{p_0-1}(x_{p_0,1}), \ldots, \phi^{p-1}(x_{p,k}), \phi^{p-2}(x_{p,k}), \ldots, \phi(x_{p,k}), x_{p,k}, \ldots, x_{p_1,d_{p_1}})\ (p_0 := \min\{p \mid d_p \neq 0\},\ \ p_1 := \max\{p \mid d_p \neq 0\})$$

are

$$([h_0]_{\mathfrak{B}},\ [\phi]_{\mathfrak{B}},\ [h_1]_{\mathfrak{B}}) = \bigoplus_{p\in\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p} \left(\varepsilon_k^{(p)}\mathscr{E}_p, \mathscr{J}_p(0), \varepsilon_k^{(p)}\mathscr{E}_p\mathscr{J}_p(0)\right).$$

In such a decomposition the number of positive and negative signs $\varepsilon_k^{(p)}$ to each p is unique and given as follows. Let $U_p := \ker\phi^{p-1} + (\ker\phi^p \cap \operatorname{range}\phi)$. *$h_0$ induces Hermitian forms $h^{(p)}$ on the quotient spaces* $\ker\phi^p/U_p$ *defined by*

$$h^{(p)}\left(x + U_p, y + U_p\right) := h_0\left(x, \phi^{p-1}(y)\right),$$

and for each $p \in \mathbb{N}_{\geqslant 1}$ with $d_p \neq 0$ the system $(x_{p,k} + U_p)_{1\leqslant k\leqslant d_p}$ is an $h^{(p)}$-orthogonal basis of $\ker\phi^p/U_p$ *such that* $h^{(p)}\left(x_{p,k} + U_p, x_{p,k} + U_p\right) = \varepsilon_k^{(p)}$. *Thus*

$$\sum_{k=1}^{d_p}\varepsilon_k^{(p)} = \text{signature}\left(h^{(p)}\right).$$

Proof. To prove the existence of the decomposition we proceed by induction. Consider the Hermitian forms

$$h_k : X \times X \to \mathbb{H}, \quad h_k(x, y) := h_0(x, \phi^k(y)).$$

If $\phi^p = 0 \neq \phi^{p-1}$, then $h_k = 0$ for all $k \geqslant p$ and, since h_0 is nondegenerate, $h_{p-1} \neq 0$. According to (5) and (6) there exists an $x \in X$ with $h_{p-1}(x, x) \in \mathbb{R}\setminus\{0\}$. Let $x_1 := x|h_{p-1}(x, x)|^{-\frac{1}{2}}$. Then

$$h_{p-1}(x_1, x_1) =: \varepsilon^{(p)} \in \{-1, 1\}.$$

Starting with x_1 define a sequence $(x_1, \ldots, x_p)$ by

$$x_k := x_{k-1} - \phi^{k-1}(x_{k-1})c_k, \qquad \text{where } c_k := \frac{\varepsilon^{(p)}}{2}h_{p-k}(x_{k-1}, x_{k-1}) \in \mathbb{R}.$$

It is then straightforward to verify, that

- $h_0(x_k, \phi^{p-1}(x_k)) = h_{p-1}(x_1, x_1) = \varepsilon^{(p)}$ for $1 \leqslant k \leqslant p$,
- $h_0(x_k, \phi^{p-j}(x_k)) = 0$, for $2 \leqslant j \leqslant k \leqslant p$,
- $h_0(x_k, \phi^{j}(x_k)) = 0$ for $1 \leqslant k \leqslant p \leqslant j$;

  in particular

$$h_0(x_p, \phi^j(x_p)) = \varepsilon^{(p)}\delta_{p-1,j} \quad \text{for all } j \in \mathbb{N}.$$

From this and the self-adjointness of ϕ, it follows that

- $h_0(\phi^{p-j}(x_p), \phi^{p-k}(x_p)) = \varepsilon^{(p)}\delta_{k+j,p+1}$,
- $h_1(\phi^{p-j}(x_p), \phi^{p-k}(x_p)) = \varepsilon^{(p)}\delta_{k+j,p+2}$ for all $k, j \in \mathbb{N}$.

Hence the Gramians of the System $\mathfrak{S} := (\phi^{p-1}(x_p), \phi^{p-2}(x_p), \ldots, \phi(x_p), x_p)$ are

$$[h_0]_{\mathfrak{S}} = \varepsilon^{(p)}\mathscr{E}_p, \quad [h_1]_{\mathfrak{S}} = \varepsilon^{(p)}\mathscr{E}_p\mathscr{J}_p(0).$$

Since $\mathscr{E}_p$ is nonsingular, we can apply the Splitting lemma and repeat the above procedure on $\langle x_p\rangle_\phi^{\perp_{h_0}}$.

The proof of the uniqueness statement is straightforward. $\square$

We present now our main result.

Theorem 3.3. *Let (h_0, ϕ, h_1) be an hs-triple on the quaternionic right vector space X.*

(a) *The quaternionic ϕ-invariant subspaces in the direct decomposition*

$$X = \bigoplus_{\lambda\in\sigma(\phi)\cap\mathbb{R}} E_\phi(\lambda) \oplus \bigoplus_{\lambda\in(\sigma(\phi)\cap\mathbb{C}^\uparrow)\setminus\mathbb{R}} \left(E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\overline{\lambda})\right) \tag{7}$$

are orthogonal to each other with respect to h_0 and h_1. h_0 is nondegenerate on each of these spaces.

(b) *Let $\lambda \in (\sigma(\phi) \cap \mathbb{C}^\uparrow) \setminus \mathbb{R}$. Then the number $d_p(\phi, \lambda)$ of Jordan blocks $\mathscr{J}_p(\lambda)$ in the Jordan canonical form of ϕ is even for each $p \in \mathbb{N}_{\geqslant 1}$.*

(c) *X has a basis $\mathfrak{B}$ such that*

$$\begin{aligned}
&([h_0]_{\mathfrak{B}}, [\phi]_{\mathfrak{B}}, [h_1]_{\mathfrak{B}}) \\
&= \bigoplus_{\lambda\in(\sigma(\phi)\cap\mathbb{C}^\uparrow)\setminus\mathbb{R}} \bigoplus_{p\in\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{\frac{1}{2}d_p(\phi,\lambda)} \left(\begin{bmatrix} 0_p & \mathscr{E}_p \\ \mathscr{E}_p & 0_p \end{bmatrix}, \begin{bmatrix} \mathscr{J}_p(\lambda) & 0_p \\ 0_p & \mathscr{J}_p(\bar{\lambda}) \end{bmatrix}, \right. \\
&\qquad\qquad \left. \begin{bmatrix} 0_p & \mathscr{E}_p\mathscr{J}_p(\bar{\lambda}) \\ \mathscr{E}_p\mathscr{J}_p(\lambda) & 0_p \end{bmatrix} \right) \\
&\oplus \bigoplus_{\lambda\in\sigma(\phi)\cap\mathbb{R}} \bigoplus_{p\in\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \left(\varepsilon_{k,\lambda}^{(p)}\mathscr{E}_p, \mathscr{J}_p(\lambda), \varepsilon_{k,\lambda}^{(p)}\mathscr{E}_p\mathscr{J}_p(\lambda) \right),
\end{aligned}$$

where $\varepsilon_{k,\lambda}^{(p)} \in \{-1, 1\}$. *Such a matrix representation is unique up to permutation of the blocks and by choice of a suitable basis* $\widetilde{\mathfrak{B}}$ *instead of* $\mathfrak{B}$ *each block of the form*

$$\left(\begin{bmatrix} 0_p & \mathscr{E}_p \\ \mathscr{E}_p & 0_p \end{bmatrix}, \begin{bmatrix} \mathscr{J}_p(\lambda) & 0_p \\ 0_p & \mathscr{J}_p(\bar{\lambda}) \end{bmatrix}, \begin{bmatrix} 0_p & \mathscr{E}_p\,\mathscr{J}_p(\bar{\lambda}) \\ \mathscr{E}_p\,\mathscr{J}_p(\lambda) & 0_p \end{bmatrix}\right)$$

can be replaced by the block

$$\left(\mathscr{E}_{2p}, \mathscr{J}_{2p}(\operatorname{Re}(\lambda), \operatorname{Im}(\lambda)), \mathscr{E}_{2p}\,\mathscr{J}_{2p}(\operatorname{Re}(\lambda), \operatorname{Im}(\lambda))\right).$$

Proof. (a) For $\lambda, \nu \in \mathbb{C}$ consider the generalized eigenspaces

$$E_\phi(\lambda) = \bigcup_{p \in \mathbb{N}_{\geqslant 1}} \ker_{\mathbb{C}}(\phi - m_\lambda)^p, \qquad E_\phi(\nu) = \bigcup_{q \in \mathbb{N}_{\geqslant 1}} \ker_{\mathbb{C}}(\phi - m_\nu)^q.$$

We use induction on $n := p + q$ to show that

$$h_0(E_\phi(\lambda), E_\phi(\nu)) = 0 \quad \text{if} \quad \{\lambda, \overline{\lambda}\} \cap \{\nu, \overline{\nu}\} = \emptyset. \tag{8}$$

For $x \in \ker_{\mathbb{C}}(\phi - m_\lambda)$, $y \in \ker_{\mathbb{C}}(\phi - m_\nu)$ we have

$$\begin{aligned} 0 &= h_0(x, (\phi - m_\nu)(y)) - h_0((\phi - m_\lambda)(x), y) \\ &= \overline{\lambda} h_0(x, y) - h_0(x, y)\nu \end{aligned}$$

and if $\mu := h_0(x, y)$ were nonzero, it followed that $\nu = \mu^{-1}\overline{\lambda}\mu$, a contradiction to Corollary 2.2. Suppose now, that $x \in \ker_{\mathbb{C}}(\phi - m_\lambda)^p$, $y \in \ker_{\mathbb{C}}(\phi - m_\lambda)^q$, where $p + q = n + 1$. Since $(\phi - m_\lambda)(x) \in \ker_{\mathbb{C}}(\phi - m_\lambda)^{p-1}$, $(\phi - m_\nu)(y) \in \ker_{\mathbb{C}}(\phi - m_\nu)^{q-1}$, the induction assumption for n yields

$$0 = h_0(x, (\phi - m_\nu)(y)) = h_0(x, \phi(y)) - h_0(x, y)\nu,$$

$$0 = h_0((\phi - m_\lambda)(x), y) = h_0(\phi(x), y) - \overline{\lambda} h_0(x, y).$$

The difference of these relations is again $0 = \overline{\lambda} h_0(x, y) - h_0(x, y)\nu$, so $h_0(x, y) = 0$. Thus we have shown that the quaternionic subspaces in the decomposition (7) are mutually orthogonal with respect to h_0. However, they are ϕ-invariant, so they are h_1-orthogonal too. Finally, if h_0 were degenerate on one of the orthogonal addends it were degenerate on X.

(b) and (c). Because of (a) we can treat the addends in (7) separately. Take $\lambda \in \sigma(\phi) \cap \mathbb{R}$ and consider the nilpotent $\mathbb{H}$-linear operator

$$\psi : E_\phi(\lambda) \to E_\phi(\lambda) \quad \psi(x) := (\phi - m_\lambda)(x).$$

According to Theorem 3.2 $E_\phi(\lambda)$ has a basis $\mathfrak{B}_\lambda$ such that

$$\begin{aligned} &([h_0]_{\mathfrak{B}_\lambda}, [\psi]_{\mathfrak{B}_\lambda}, [h_0(\psi(\cdot), \cdot)]_{\mathfrak{B}_\lambda}) \\ &\quad = \bigoplus_{p \in \mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \left(\varepsilon_{k,\lambda}^{(p)} \mathscr{E}_p, \mathscr{J}_p(0), \varepsilon_{k,\lambda}^{(p)} \mathscr{E}_p \mathscr{J}_p(0)\right), \end{aligned}$$

with unique $\varepsilon_{k,\lambda}^{(p)} \in \{-1, 1\}$. Since $\phi|_{E_\phi(\lambda)} = m_\lambda + \psi$, we have

$$([h_0]_{\mathfrak{B}_\lambda}, [\phi]_{\mathfrak{B}_\lambda}, [h_1]_{\mathfrak{B}_\lambda}) = \bigoplus_{p\in\mathbb{N}_{\geqslant 1}} \bigoplus_{k=1}^{d_p(\phi,\lambda)} \left(\varepsilon_{k,\lambda}^{(p)}\mathscr{E}_p,\ \mathscr{J}_p(\lambda),\ \varepsilon_{k,\lambda}^{(p)}\mathscr{E}_p\mathscr{J}_p(\lambda)\right).$$

Next take a $\lambda \in (\sigma(\phi) \cap \mathbb{C}^{\uparrow}) \setminus \mathbb{R}$. We need the following:

Lemma 3.4. *The mapping*

$$(x, y) \overset{s}{\longmapsto} s(x, y) := h_0(x\mathrm{J}, y), \quad x, y \in E_\phi(\lambda)$$

is a $\mathbb{C}$-valued and nondegenerate skew-symmetric form on $E_\phi(\lambda)$. The $\mathbb{C}$-linear operator $\psi := (\phi - m_\lambda)|_{E_\phi(\lambda)}$ is self-adjoint with respect to s, i.e.

$$s(\psi(x), y) = s(x, \psi(y)) \quad \textit{for all } x, y \in E_\phi(\lambda).$$

Taking this for granted for the moment we see, that the mappings

$$s_j : E_\phi(\lambda) \times E_\phi(\lambda) \to \mathbb{C}, \quad s_j(x, y) := s(x, \psi^j(y)), \quad j \in \mathbb{N},$$

are skew-symmetric forms. Since ψ is nilpotent of degree p, say, we have $s_j = 0$ for $j \geqslant p$ and there exists a $y \in E_\phi(\lambda)$ with $\psi^{p-1}(y) \neq 0$. Moreover, since s is nondegenerate, there must be an $x_1 \in E_\phi(\lambda)$ such that $s_{p-1}(x_1, y) = s(x_1, \psi^{p-1}(y)) = 1$. Starting with x_1 define a sequence $(x_1, \ldots, x_p)$ by

$$x_k := x_{k-1} - \psi^{p-k}(x_{k-1})c_k, \quad \text{where} \quad c_k := s_{p-k}(x_{k-1}, y) \in \mathbb{C}.$$

Then it holds for $x := x_p$ that

$$s_j(x, y) = s(x, \psi^j(y)) = \delta_{p-1,j} \quad \text{for all } j \in \mathbb{N}. \tag{9}$$

Now define for $k = 1, \ldots, p$,

$$\begin{aligned} &u_k := \psi^{p-k}(x)\mathrm{J}, \quad v_k := \psi^{p-k}(y), \\ &\xi_k := u_k\left(\tfrac{1+\mathrm{i}}{2}\right) + v_k\left(\tfrac{1-\mathrm{i}}{2}\right), \quad \eta_k := u_k\left(\tfrac{1-\mathrm{i}}{2}\right) + v_k\left(\tfrac{1+\mathrm{i}}{2}\right). \end{aligned}$$

Then we have for $k = 1, \ldots, p$,

$$u_k = \xi_k\left(\tfrac{1-\mathrm{i}}{2}\right) + \eta_k\left(\tfrac{1+\mathrm{i}}{2}\right), \quad v_k = \xi_k\left(\tfrac{1+\mathrm{i}}{2}\right) + \eta_k\left(\tfrac{1-\mathrm{i}}{2}\right)$$

and

$$\begin{aligned} &\phi(u_k) = u_k\overline{\lambda} + u_{k-1}, \qquad \phi(\xi_k) = \xi_k \operatorname{Re}(\lambda) + \eta_k \operatorname{Im}(\lambda) + \xi_{k-1}, \\ &\phi(v_k) = v_k\lambda + v_{k-1}, \qquad \phi(\eta_k) = \xi_k(-\operatorname{Im}(\lambda)) + \eta_k \operatorname{Re}(\lambda) + \eta_{k-1}, \end{aligned} \tag{10}$$

where $u_0 := v_0 := \xi_0 := \eta_0 := 0$. Hence the systems $\mathfrak{S} := (u_1, \ldots, u_p, v_1, \ldots, v_p)$ and $\widetilde{\mathfrak{S}} := (\xi_1, \eta_1, \ldots, \xi_p, \eta_p)$ span the same ϕ-invariant subspace$U$. From (9) and the fact, that $s_j(z,z)=0$ for all $j\in\mathbb{N}$, $z\in E_\phi(\lambda)$, we obtain

$$\begin{aligned}
h_0(u_k,v_j) &= s(\psi^{p-k}(x),\psi^{p-j}(y)) = s(x,\psi^{2p-k-j}(y)) = \delta_{p+1,k+j},\\
h_0(u_k,u_j) &= s(\psi^{p-k}(x),\psi^{p-j}(x))\mathrm{J} = s(x,\psi^{2p-k-j}(x))\mathrm{J}\\
&= s_{2p-k-j}(x,x)\mathrm{J} = 0,\\
h_0(v_k,v_j) &= \mathrm{J}s_{2p-k-j}(y,y) = 0,\\
h_0(\xi_k,\eta_j) &= \left(\frac{1-\mathrm{i}}{2}\right)h_0(u_k,u_j)\left(\frac{1-\mathrm{i}}{2}\right)+\left(\frac{1+\mathrm{i}}{2}\right)h_0(v_k,v_j)\left(\frac{1+\mathrm{i}}{2}\right)\\
&\quad+\left(\frac{1-\mathrm{i}}{2}\right)h_0(u_k,v_j)\left(\frac{1+\mathrm{i}}{2}\right)+\left(\frac{1+\mathrm{i}}{2}\right)h_0(v_k,u_j)\left(\frac{1-\mathrm{i}}{2}\right)\\
&= \delta_{p+1,k+j},\\
h_0(\xi_k,\xi_j) &= \left(\frac{1-\mathrm{i}}{2}\right)h_0(u_k,u_j)\left(\frac{1+\mathrm{i}}{2}\right)+\left(\frac{1+\mathrm{i}}{2}\right)h_0(v_k,v_j)\left(\frac{1-\mathrm{i}}{2}\right)\\
&\quad+\left(\frac{1-\mathrm{i}}{2}\right)h_0(u_k,v_j)\left(\frac{1-\mathrm{i}}{2}\right)+\left(\frac{1+\mathrm{i}}{2}\right)h_0(v_k,u_j)\left(\frac{1+\mathrm{i}}{2}\right)\\
&= 0,\\
h_0(\eta_k,\eta_j) &= \left(\frac{1+\mathrm{i}}{2}\right)h_0(u_k,u_j)\left(\frac{1-\mathrm{i}}{2}\right)+\left(\frac{1-\mathrm{i}}{2}\right)h_0(v_k,v_j)\left(\frac{1+\mathrm{i}}{2}\right)\\
&\quad+\left(\frac{1+\mathrm{i}}{2}\right)h_0(u_k,v_j)\left(\frac{1+\mathrm{i}}{2}\right)+\left(\frac{1-\mathrm{i}}{2}\right)h_0(v_k,u_j)\left(\frac{1-\mathrm{i}}{2}\right)\\
&= 0.
\end{aligned}$$

Hence the h_0-Gramians of the Systems $\mathfrak{S}$ and $\tilde{\mathfrak{S}}$ are

$$[h_0]_{\mathfrak{S}} = [h_0]_{\tilde{\mathfrak{S}}} = \begin{bmatrix} 0_p & \mathscr{E}_p \\ \mathscr{E}_p & 0_p \end{bmatrix} = \mathscr{E}_{2p}.$$

Since $\mathscr{E}_{2p}$ has full rank, $\mathfrak{S}$ and $\tilde{\mathfrak{S}}$ are linearly independent systems. Thus $\mathfrak{S}$ and $\tilde{\mathfrak{S}}$ are bases of U, and by (10) we have

$$[\phi]_{\mathfrak{S}} = \begin{bmatrix} \mathscr{J}_p(\bar{\lambda}) & 0_p \\ 0_p & \mathscr{J}_p(\lambda) \end{bmatrix}, \quad [\phi]_{\tilde{\mathfrak{S}}} = \mathscr{J}_{2p}(\mathrm{Re}(\lambda),\mathrm{Im}(\lambda)).$$

Apply now the Splitting lemma and repeat the procedure on $U^{\perp_{h_0}}$.

It remains to prove the lemma about the properties of s. We have for all $x,y\in X$, $\lambda\in\mathbb{C}$:

$$\begin{aligned}
s(x\lambda,y) &= h_0(x\lambda\mathrm{J},y) = h_0(x\mathrm{J}\bar{\lambda},y) = \lambda h_0(x\mathrm{J},y) = \lambda s(x,y),\\
s(x,y\lambda) &= h_0(x\mathrm{J},y\lambda) = h_0(x\mathrm{J},y)\lambda = s(x,y)\lambda.
\end{aligned}$$

Thus s is $\mathbb{C}$-bilinear. To see that $s(x,y)\in\mathbb{C}$ for all $x\in\ker_{\mathbb{C}}(\phi-m_\lambda)^p$, $y\in\ker_{\mathbb{C}}(\phi-m_\lambda)^q$ and all $p,q\in\mathbb{N}_{\geqslant 1}$, we use induction on $n:=p+q$ again. For $x,y\in\ker_{\mathbb{C}}(\phi-m_\lambda)$ we have

$$\begin{aligned} 0 &= h_0((\phi - m_\lambda)(x)\mathrm{j}, y) - h_0(x\mathrm{j}, (\phi - m_\lambda)(y)) \\ &= s(x, y)\lambda - \lambda s(x, y). \end{aligned}$$

So $\lambda \in \mathbb{C} \setminus \mathbb{R}$ commutes with $s(x, y)$, which is merely possible if $s(x, y) \in \mathbb{C}$, as easily seen. Suppose now, that $x \in \ker_{\mathbb{C}}(\phi - m_\lambda)^p$, $y \in \ker_{\mathbb{C}}(\phi - m_\lambda)^q$, where $p + q = n + 1$. The induction assumption for n yields

$$0 = s(x, (\phi - m_\lambda)(y)) = h_0(x\mathrm{j}, \phi(y)) - s(x, y)\lambda,$$

$$0 = s((\phi - m_\lambda)(x), y) = h_0(\phi(x)\mathrm{j}, y) - \lambda s(x, y).$$

Take the difference of these relations to see that λ commutes with $s(x, y)$.

To prove the skew-symmetry of s consider the equation

$$s(x, y) + s(y, x) = h(x\mathrm{j}, y) + h(y\mathrm{j}, x) = -2\mathrm{j}\mathrm{Re}(h(x, y)).$$

Since s is $\mathbb{C}$-valued, it follows that $s(x, y) + s(y, x) = 0$.

We show now that s is nondegenerate. According to (a) h_0 is nondegenerate on $E_\phi(\lambda) \oplus_{\mathbb{C}} E_\phi(\overline{\lambda})$, where $E_\phi(\overline{\lambda}) = E_\phi(\lambda)\mathrm{j}$. Hence to each $y \in E_\phi(\lambda) \setminus \{0\}$ there exist $x_1, x_2 \in E_\phi(\lambda)$ such that

$$0 \neq h_0(x_1 + x_2\mathrm{j}, y) = \mathrm{j}s(x_1, y) + s(x_2, y),$$

i.e. $s(x_1, y) \neq 0$ or $s(x_2, y) \neq 0$. The s-self-adjointness of ψ is obvious. $\square$

It is a basic result that two Hermitian forms over $\mathbb{C}$ can be simultaneously diagonalized if one form is positive definite. This theorem also holds in the quaternionic case.

Corollary 3.5. *Let X be an n-dimensional quaternionic right vector space and (h_0, ϕ, h_1) be an hs-triple on X, where h_0 is positive definite (i.e. $h_0(x, x) > 0$ for all $x \in X \setminus \{0\}$). Then $\{\lambda_1, \ldots, \lambda_n\} := \sigma(\phi) \subseteq \mathbb{R}$ and there is a basis $\mathfrak{B}$ of X such that*

$$([h_0]_{\mathfrak{B}}, [\phi]_{\mathfrak{B}}, [h_1]_{\mathfrak{B}}) = (I_n, \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \mathrm{diag}(\lambda_1, \ldots, \lambda_n)),$$

where I_n is the $n \times n$ unit matrix.

Proof. If ϕ had a nonreal eigenvalue, then the diagonal of $[h_0]_{\mathfrak{B}}$ in the canonical matrix representation would contain at least one zero, which corresponds to an $x \in X \setminus \{0\}$ with $h_0(x, x) = 0$. So h_0 were not positive definite. Similar arguments hold if $[h_0]_{\mathfrak{B}}$ contains a block $\varepsilon_{k,\lambda}^{(p)} \mathscr{E}_p$ with $p \geqslant 2$ or $\varepsilon_{k,\lambda}^{(p)} = -1$. $\square$

References

[1] S. Adler, Quaternionic Quantum Mechanics and Quantum Fields, Oxford University Press, Oxford, 1995.

[2] A. Baker, Right eigenvalues for quaternionic matrices: A topological approach, Linear Algebra Appl. 286 (1999) 303–309.
[3] E. Brieskorn, Lineare Algebra und Analytische Geometrie, Band 2, Vieweg-Verlag, 1993.
[4] P.M. Cohn, The similarity reduction of matrices over a skew field, Math. Z. 132 (1973) 151–163.
[5] I. Gohberg, P. Lancaster, L. Rodman, Matrices and Indefinite Scalar Products, Birkhuser, 1983.
[6] N. Jacobson, Normal semilinear transformations, Amer. J. Math. 61 (1939) 45–58.
[7] N. Jacobson, Theory of Rings, Math. Surveys Nr. 2, Am. Math. Soc. (1943).
[8] M. Koecher and R. Remmert, Hamilton's Quaternions, in Numbers, Graduate Texts in Mathematics 123, Springer, Berlin, 1991.
[9] H.C. Lee, Eigenvalues and canonical forms of matrices with quaternion coefficients, Proc. Royal Irish Academy, vol. 52, Sect. A (1949) 253–260.
[10] A.I. Mal'cev, Foundations of Linear Algebra, Freeman, San Francisco, 1963.
[11] V.V. Sergeichuk, Classification of sesquilinear forms, pairs of Hermitian forms, self-conjugate and isometric operators over the division ring of quaternions, Math. Notes 49 (4) (1991) 409–414.
[12] R.C. Thompson, Pencils of complex and real symmetric and skew matrices, Linear Algebra Appl. 147 (1991) 323–371.
[13] W.C. Waterhouse, Pairs of quadratic forms, Invent. Math. 37:157–164 (1976).
[14] L.A. Wolf, Similarity of matrices in which the elements are real quaternions, Bull. Amer. Math. Soc. 42 (1936) 737–743.
[15] R.M.W. Wood, Quaternionic eigenvalues, Bull. London Math. Soc. 42 (1985) 137–138.
[16] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (1997) 21–57.